

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 310.315 Seconds
(without alignments)
9133.322 Million cell updates/sec

Title: US-10-010-408-1_COPY_1534_1708
Perfect score: 175
Sequence: 1 AGTCCAGGAATTGAGCTT.....GCCTAGATAACACCCAAA 175

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_man:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 173.4 | 99.1 | 762 | 14 | BQ195526 |
| 2 | 126.2 | 72.1 | 418 | 9 | AI225477 |
| 3 | 124.6 | 71.2 | 313 | 10 | BB367824 |
| 4 | 123 | 70.3 | 315 | 10 | BB374499 |
| 5 | 119.8 | 68.5 | 307 | 10 | BB220676 |
| 6 | 119.8 | 68.5 | 337 | 10 | BB319151 |

| | | | | | |
|----|-------|------|------|----|----------|
| 7 | 119.8 | 68.5 | 345 | 10 | BB222795 |
| 8 | 113.4 | 64.8 | 281 | 10 | BB309266 |
| 9 | 113.4 | 64.8 | 316 | 10 | BB518921 |
| 10 | 110.2 | 63.0 | 218 | 10 | BB213539 |
| 11 | 110.2 | 63.0 | 289 | 10 | BB186801 |
| 12 | 110.2 | 63.0 | 290 | 10 | BB515213 |
| 13 | 104.8 | 59.9 | 369 | 10 | BB925244 |
| 14 | 100.2 | 57.3 | 324 | 10 | BB580051 |
| 15 | 74.4 | 42.5 | 792 | 12 | BF138093 |
| 16 | 33.6 | 19.2 | 675 | 17 | AZ574808 |
| 17 | 33.6 | 19.2 | 712 | 13 | BI147316 |
| 18 | 33.4 | 19.1 | 360 | 17 | AQ345221 |
| 19 | 33.4 | 19.1 | 506 | 10 | BB701454 |
| 20 | 33.2 | 19.0 | 469 | 9 | AL800293 |
| 21 | 32.2 | 18.4 | 317 | 12 | BF077992 |
| 22 | 32.2 | 18.4 | 441 | 10 | BB775647 |
| 23 | 32.2 | 18.4 | 589 | 12 | BF712425 |
| 24 | 32.2 | 18.4 | 620 | 10 | AV751782 |
| 25 | 32 | 18.3 | 677 | 17 | AZ574825 |
| 26 | 32 | 18.3 | 682 | 17 | AZ574814 |
| 27 | 32 | 18.3 | 744 | 17 | CNS04PR1 |
| 28 | 32 | 18.3 | 1089 | 17 | CNS031YH |
| 29 | 31.8 | 18.2 | 790 | 9 | AL820075 |
| 30 | 31.6 | 18.1 | 799 | 13 | BI757228 |
| 31 | 31.4 | 17.9 | 1135 | 14 | BQ278875 |
| 32 | 31.2 | 17.8 | 412 | 10 | AW694655 |
| 33 | 31.2 | 17.8 | 454 | 9 | AL379783 |
| 34 | 31 | 17.7 | 263 | 9 | AV063522 |
| 35 | 31 | 17.7 | 327 | 12 | BF549593 |
| 36 | 31 | 17.7 | 373 | 17 | AQ120776 |
| 37 | 31 | 17.7 | 434 | 13 | BM256378 |
| 38 | 31 | 17.7 | 506 | 10 | BE485090 |
| 39 | 31 | 17.7 | 639 | 13 | BJ254458 |
| 40 | 31 | 17.7 | 691 | 17 | AQ752684 |
| 41 | 31 | 17.7 | 697 | 17 | AZ127785 |
| 42 | 31 | 17.7 | 740 | 14 | BQ210400 |
| 43 | 31 | 17.7 | 799 | 17 | AZ185214 |
| 44 | 31 | 17.7 | 929 | 14 | BQ425393 |
| 45 | 30.8 | 17.6 | 418 | 14 | W97990 |

ALIGNMENTS

RESULT 1
BQ195526/c
LOCUS
DEFINITION
UI-R-CNI-cmq-k-07-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone
BQ195526
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 762)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A

APPENDIX: BONAUDO ET AL Y2

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab Clone distribution: Clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..762
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cmq-k-07-0-UI"
/clone_lib="UI-R-CNI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Urogen Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AMV through R-CA0-AXS, R-CA0-AXZ through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOP through R-CA0-BPJ, R-CA0-BPA through R-CA0-BPG, R-CA0-BBA through R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR, R-CA1-BJT through R-CA1-BKE, R-CA1-BKD, R-CA1-BKF, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLU-V, R-CA1-BNR, and R-CA1-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCM, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVN, R-CT0-BPV through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BFC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BVOP (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BVOP-AOI through R-BVOP-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CV0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZI through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLW-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkz-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with NotI and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG LIB=UI-R-CNI
TAG TISSUE=cervix
TAG_SEQ=GACCA"

BASE COUNT 177 a 176 c 198 g 211 t
ORIGIN

Query Match 99.1%; Score 173.4; DB 14; Length 762;
Best Local Similarity 99.4%; Pred. No. 2.6e-48;
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCCAGGAAGTGGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
Db 202 AGTCCAGGAAGTGGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 143
Qy 61 AGGAAGGCTCCACACCTCTGCGAGGCCGCTTCTTCAGCATGAGAAACAGG 120
Db 142 AGGAGGCTCCACACCTCTGCGAGGCCGCTTCTTCAGCATGAGAAACAGG 83
Qy 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGATAAACACCCAAA 175
Db 82 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGATAAACACCCAAA 28

RESULT 2

AI225477 us88b01.yl Soares_NMPu 418 bp mRNA linear EST 29-OCT-1998
LOCUS
DEFINITION mRNA sequence.

AI225477 GI:3808530

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 418)

REFERENCE

AUTHORS

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Weising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

The WashU-HMI Mouse EST Project

JOURNAL

Unpublished (1996)

A7PENDX BONDADO ET AL 42